Casey Grage and Renee Zha

Checkpoint #2: Question 1

Data integration questions:

1. **Are the identity groups most represented in plaintiffs from the settlement data also most likely to be victims in the complaint reports?**

For our first two questions, we intended to combine demographic data about civilians from the settlement and cpdb datasets in a pipeline similar to the following:

- **(Make table of total no. of plaintiffs in each identity group)** Identify identity groups of plaintiffs in the settlement data using demographic information from cpdb data (make a table joining the names from plaintiffs in cpdb to the complainants' identity groups, if possible)
- **(Make table of no. of victims in each identity group)** Identify most common identity groups of victims from cpdb (done in checkpoint #1)

We were unable to find sufficient identifying or demographic information about the plaintiffs in the provided core settlement database that was published on Canvas, so we had to improvise methods of collecting this data…

We mapped settlement location to the corresponding cpdb beat_id. We then found the majority race of each beat_id, and assigned that race to the race of the plaintiff. This allowed us to find the percentage of each race in total settlements. We also looked up the median income of each beat_id and used that data to look at trends between race, income_area, and no. of settlements. This method also normalizes the number of allegations from people of a given identity group in a neighborhood by their representation in that area's population. Otherwise it is difficult to distinguish whether one group is represented among the plaintiffs.

Settlement provides lat, lon, location address, settlement ID

Plaintiff provides name, and connects to settlement_plaintiff which has a settlement ID

FIRST (find most represented races in plaintiffs)

- Cpdb data_area provides median_income, polygon, and data_racepopulation
- Map settlement lat/lon to cpdb polygons

```
CREATE TABLE mapped_settlements AS SELECT a.id,c.case_no FROM
data_area a, cases_case c WHERE ST_INTERSECTS(a.polygon,
ST_SetSRID(ST_MakePoint(c.lon,c.lat),4326)) AND a.id NOTNULL ORDER BY
a.id desc;
```

(`SELECT a.description,c.case_no FROM data_area a, cases_case c WHERE ST_INTERSECTS(a.polygon, ST_SetSRID(ST_MakePoint(c.lon,c.lat),4326)) AND a.description NOTNULL;` for the names of the neighborhoods. Just interesting!)

- Make table of majority race per area_id

```
CREATE TABLE mapped_majorityrace AS SELECT rp.* FROM
data_racepopulation rp INNER JOIN (SELECT area_id, MAX(count) AS
MajorityRace FROM data_racepopulation GROUP BY area_id) groupedrp ON
rp.area_id = groupedrp.area_id AND rp.count = groupedrp.MajorityRace
ORDER BY rp.area_id desc;

CREATE TABLE majorityrace_sansid AS SELECT race, count, area_id FROM
mapped_majorityrace GROUP BY race, count, area_id ORDER BY area_id
desc;
```

Note: the area_id's for data_racepopulation only start at 426. I guess the area_id's less than that don't have corresponding race data in the database.

- Assign majority race (from data_racepopulation of the polygon) to the race of that settlement

```
CREATE TABLE settlement_races AS SELECT mapped_settlements.*,
majorityrace_sansid.* FROM mapped_settlements JOIN
majorityrace_sansid ON majorityrace_sansid.area_id =
mapped_settlements.id;
```

- THEN find percentage of each race in total settlements

```
SELECT race, COUNT(distinct case_no) FROM settlement_races GROUP BY
race ORDER BY COUNT(*) DESC;
```

```
   race    | count
-----------+-------
 Black     |   462
 Hispanic  |   256
 White     |   179
 Asian     |     8
(4 rows)
```

Total: 905

Black: 462/905 * 100 = **51.0497%**

Hispanic: 256/905 * 100 = **28.2873%**

White: 179/905 * 100 = **19.7790%**

Asian: 8/905 * 100 = **0.883978%**

Native American: 0/905 = **0%**

SECOND (find most represented races in complainants)

- Make table of total no. of complaints each identity group has filed

  QUERY (For officer *and* civilian complainants):

```
SELECT race, COUNT(distinct allegation_id) FROM
data_complainant GROUP BY race ORDER BY COUNT(*) DESC;
              race               | count
---------------------------------+-------
 Black                           | 33032
 White                           | 11591
 Hispanic                        | 10220
 Asian/Pacific Islander          |   485
 Native American/Alaskan Native |    74
(5 rows)
```

Total: 55402

Black: 33032/55402 * 100 = **59.6223%**

Hispanic: 10220/55402 * 100 = **18.4470%**

White: 11591/55402 * 100 = **20.9216%**

Asian: 485/55402 * 100 = **0.875419%**

Native American: 74/55402 * 100 = **0.133569%**

Compare Complainant Races to Settlement Plaintiff Races:

Race	***Complainant***	***Plaintiff***	***% Change from Complainant to Plaintiff***
Black	**51.0497%**	**59.6223%**	**+ 8.5726%**
Hispanic	**28.2873%**	**18.4470%**	**- 9.8403**
White	**19.7790%**	**20.9216%**	**+ 1.1426**
Asian	**0.883978%**	**0.875419%**	**- 0.008559**
Native American	**0%**	**0.133569%**	**+ 0.133569%**

What we learned: There doesn't seem to be a huge difference between complainant racial breakdown and plaintiff racial breakdown. The main notable change was that the percentage of black plaintiffs is higher than the percentage of black complainants, while the opposite is true of hispanics. It is possible that black civilians are more likely to file more serious allegations than other races. It is also possible that hispanics are less likely to follow up complaints with legal action. This could be due to obstacles against them (lawyers are less likely to take their cases, etc.).

THIRD

- *We found median_income to be very lacking??? There's just a ton of beat_id's that have no corresponding median_income.*

`SELECT* median_income FROM data_area;` shows median_income is missing a lot of data….

- Join settlement_races to neighborhood of that area_id

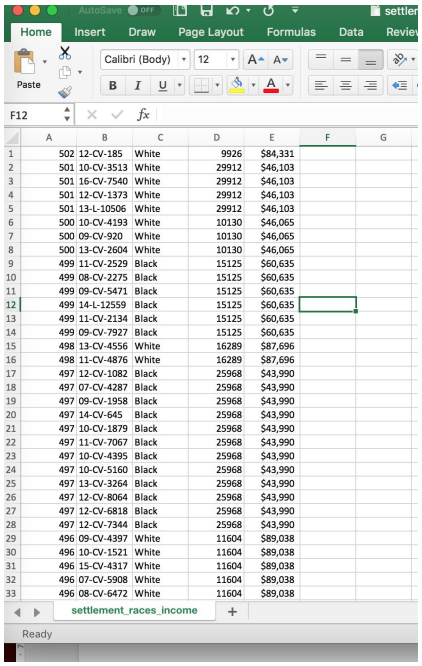
```
CREATE TABLE settlement_races_income AS SELECT
settlement_races.area_id, settlement_races.case_no,
settlement_races.race, settlement_races.count,
data_area.median_income FROM settlement_races JOIN data_area ON
settlement_races.area_id = data_area.id;

SELECT* FROM settlement_races_income WHERE median_income NOTNULL;
```

I've decided there is enough median_income data. No need to look for external data (external data wouldn't even necessarily match up to area_id)

- Export table to CSV

```
\copy settlement_races_income to
'/Users/caseygrage/Downloads/settlement_races_income.csv' WITH
(FORMAT csv);
```



	A	B	C	D	E
1	502	12-CV-185	White	9926	$84,331
2	501	10-CV-3513	White	29912	$46,103
3	501	16-CV-7540	White	29912	$46,103
4	501	12-CV-1373	White	29912	$46,103
5	501	13-L-10506	White	29912	$46,103
6	500	10-CV-4193	White	10130	$46,065
7	500	09-CV-920	White	10130	$46,065
8	500	13-CV-2604	White	10130	$46,065
9	499	11-CV-2529	Black	15125	$60,635
10	499	08-CV-2275	Black	15125	$60,635
11	499	09-CV-5471	Black	15125	$60,635
12	499	14-L-12559	Black	15125	$60,635
13	499	11-CV-2134	Black	15125	$60,635
14	499	09-CV-7927	Black	15125	$60,635
15	498	13-CV-4556	White	16289	$87,696
16	498	11-CV-4876	White	16289	$87,696
17	497	12-CV-1082	Black	25968	$43,990
18	497	07-CV-4287	Black	25968	$43,990
19	497	09-CV-1958	Black	25968	$43,990
20	497	14-CV-645	Black	25968	$43,990
21	497	10-CV-1879	Black	25968	$43,990
22	497	11-CV-7067	Black	25968	$43,990
23	497	10-CV-4395	Black	25968	$43,990
24	497	10-CV-5160	Black	25968	$43,990
25	497	13-CV-3264	Black	25968	$43,990
26	497	12-CV-8064	Black	25968	$43,990
27	497	12-CV-6818	Black	25968	$43,990
28	497	12-CV-7344	Black	25968	$43,990
29	496	09-CV-4397	White	11604	$89,038
30	496	10-CV-1521	White	11604	$89,038
31	496	15-CV-4317	White	11604	$89,038
32	496	07-CV-5908	White	11604	$89,038
33	496	08-CV-6472	White	11604	$89,038

- Reformat strings of median_income to ints

```
CREATE TABLE races_income2 (area_id varchar, case_no varchar, count
varchar, string_median_income varchar, race varchar, median_income
int);

\copy races_income2 FROM
'/Users/caseygrage/Downloads/races_income2.csv' WITH (FORMAT csv);
```

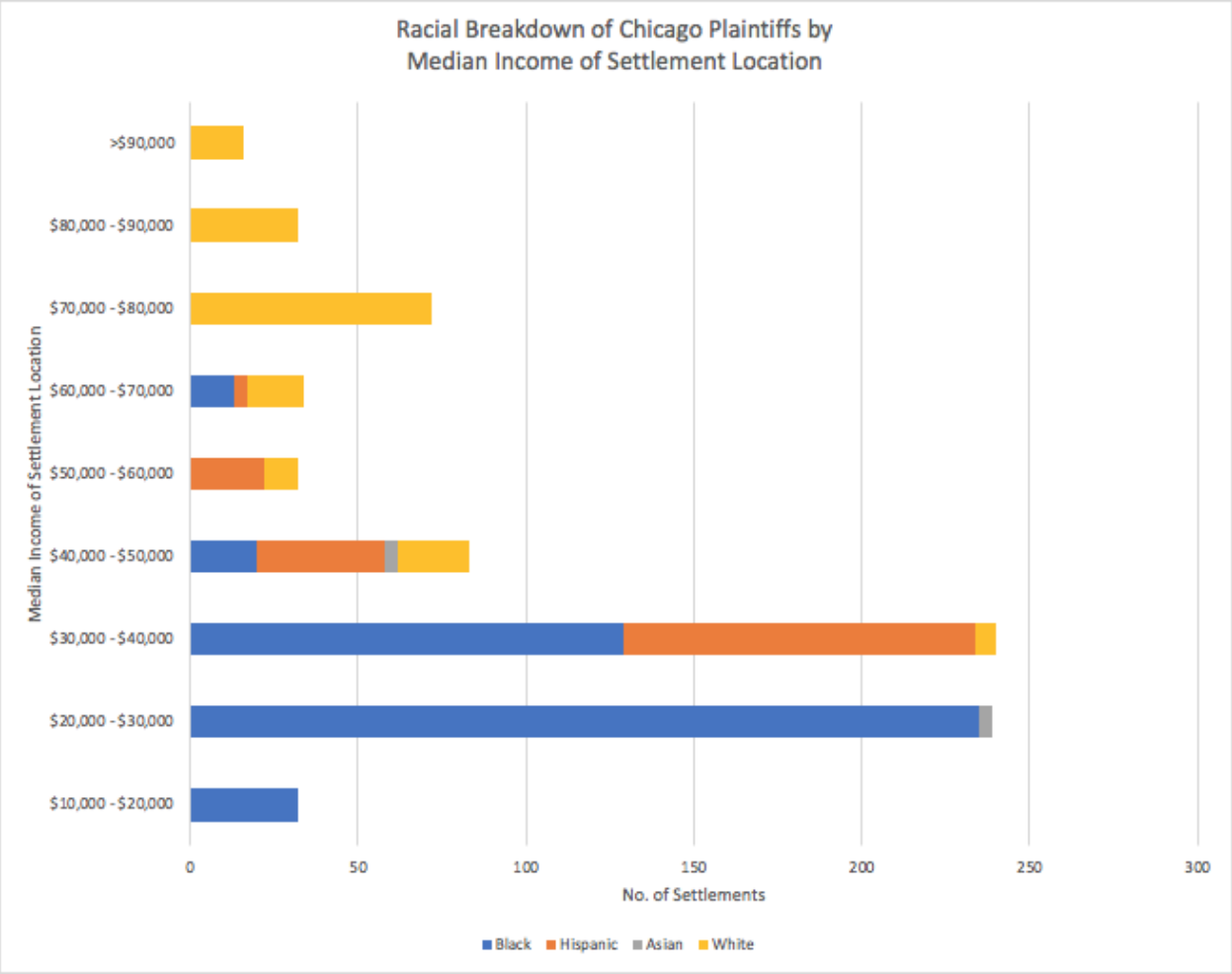
- Find number of complainants per income group (by race)

```
SELECT race, COUNT(*) FROM races_income2 WHERE median_income < 10000
GROUP BY race ORDER BY COUNT(*);
race | count
------+-------
(0 rows)
SELECT race, COUNT(*) FROM races_income2 WHERE median_income < 20000
GROUP BY race ORDER BY COUNT(*);
 race  | count
-------+-------
 Black |    32
(1 row)
SELECT race, COUNT(*) FROM races_income2 WHERE (median_income < 30000
AND median_income > 20000) GROUP BY race ORDER BY COUNT(*);
 race  | count
-------+-------
 Asian |     4
 Black |   235
(2 rows)
SELECT race, COUNT(*) FROM races_income2 WHERE (median_income < 40000
AND median_income > 30000) GROUP BY race ORDER BY COUNT(*);
   race   | count
----------+-------
 White    |     6
 Hispanic |   105
 Black    |   129
(3 rows)
SELECT race, COUNT(*) FROM races_income2 WHERE (median_income < 50000
AND median_income > 40000) GROUP BY race ORDER BY COUNT(*);
   race   | count
----------+-------
 Asian    |     4
 Black    |    20
 White    |    21
 Hispanic |    38
(4 rows)
```

```
SELECT race, COUNT(*) FROM races_income2 WHERE (median_income < 60000
AND median_income > 50000) GROUP BY race ORDER BY COUNT(*);
   race   | count
----------+-------
 White    |    10
 Hispanic |    22
(2 rows)
SELECT race, COUNT(*) FROM races_income2 WHERE (median_income < 70000
AND median_income > 60000) GROUP BY race ORDER BY COUNT(*);
   race   | count
----------+-------
 Hispanic |     4
 Black    |    13
 White    |    17
(3 rows)
SELECT race, COUNT(*) FROM races_income2 WHERE (median_income < 80000
AND median_income > 70000) GROUP BY race ORDER BY COUNT(*);
 race  | count
-------+-------
 White |    72
(1 row)
SELECT race, COUNT(*) FROM races_income2 WHERE (median_income < 90000
AND median_income > 80000) GROUP BY race ORDER BY COUNT(*);
 race  | count
-------+-------
 White |    32
(1 row)
SELECT race, COUNT(*) FROM races_income2 WHERE (median_income <
100000 AND median_income > 90000) GROUP BY race ORDER BY COUNT(*);
 race  | count
-------+-------
 White |    15
(1 row)
SELECT race, COUNT(*) FROM races_income2 WHERE (median_income >=
100000) GROUP BY race ORDER BY COUNT(*);
 race  | count
-------+-------
```

```
 White |      1
(1 row)
```

Put that data into excel to make a nice stacked bar chart!



What we learned: The majority of settlements happen in lower income neighborhoods. The racial breakdown of these neighborhoods' populations is such that black plaintiffs are by far the most common in low income neighborhoods, while white plaintiffs are the most common in wealthier neighborhoods. The majority of hispanic plaintiffs have settlements that occur in middle to low income areas. It is also interesting to see that the only income area in which all four major races are represented is the lower middle class $40,000 - $50,000. This graph shows that plaintiffs are largely lower income, and largely black and hispanic. However, the number of plaintiffs steadies for the middle to upper classes.

2. Are certain identity groups overrepresented in the complaints and settlement database compared to the general population?

Because of the problems we ran into in #1, we were unable to analyze the overall demographics of *civilians* (plaintiffs) in both the settlement and cpdb dataset. However, to gain some initial insight into this problem, we found the percent chance that any given race will be a complainant (by comparing number of complainants by race compared to their total Chicago population - using the data_racepopulation table from cpdb).

Make table of total no. of complaints each identity group has filed

QUERY (For officer *and* civilian complainants):

```
SELECT race, COUNT(distinct allegation_id) FROM
data_complainant GROUP BY race ORDER BY COUNT(*) DESC;
```

```
               race               | count
----------------------------------+-------
 Black                            | 33032
 White                            | 11591
 Hispanic                         |  5339
                                  |  4881
 Asian/Pacific Islander           |   485
 Native American/Alaskan Native   |    74
(6 rows)
```

Compare to general population of identity group

QUERY:

```
SELECT race, SUM(count) AS "Total Counts" FROM
data_racepopulation GROUP BY race;
```

```
      race        | Total Counts
------------------+--------------
 Black            |      1712474
 Hispanic         |      1569483
 Native American  |         4097
 Other            |        91342
 White            |      1729593
 Asian            |       305590
(6 rows)
```

Other: 0.05343653522 * 100 = **5.344%**

Black: 0.01928905198 * 100 = **1.929%**

Native American: 0.01806199658 * 100 = **1.806%**

White: 0.006701576614 * 100 = **0.6701%**

Hispanic: 0.003401757139 * 100 = **0.3401%**

Asian: 0.001587093819 * 100 = **0.1587%**

From this initial exploration, we see that there is a wide range in the percentage of people in a certain identity group who file complaints. In particular, the rate of complaints from people who are black is almost triple that of those who are white, and 12x the rate of those who are Asian. This suggests that there may be underlying factors (cultural, societal, biases, etc.) that affect these rates.

3. Do police officers of certain identity groups have more allegations against them in low income areas than in higher income areas?

We used Trifacta Wrangler to join the data_officerallegations table with information from data_officer, data_allegation, and data_area, in order to combine demographic information about the listed officers with the areas (beats) that the allegations took place.

Used https://www.chicagocomputerclasses.com/average-city-chicago-income/ to find the neighborhoods with the highest and lowest average incomes.

Used https://data.cpdp.co/data/bgKwNB/ to find the beats within neighborhoods of interest.

Used www.areavibes.com to find racial demographics of neighborhoods of interest.

Queries to import CSV to Postgres:

```
CREATE TABLE data_oa_with_area (beat_id varchar, allegation_id
varchar, race varchar, name varchar);
```

```
\copy data_oa_with_area FROM
'/Users/caseygrage/Downloads/data_oa_with_area.csv' WITH (FORMAT
csv);
```

Higher income neighborhoods (descending order):

1) Near North Side: Avg Income $88,669.00
    a) Table of number of allegations per identity group in this neighborhood

```
SELECT race, COUNT(*) FROM data_oa_with_area WHERE name='1833' OR
name='1834' OR name='1832' OR name='1831' OR name='1824' OR
name='1823' OR name='1822' OR name='1821' GROUP BY race ORDER BY
COUNT(*) ASC;
```

```
              race               | count
---------------------------------+-------
 Native American/Alaskan Native |     5
 Asian/Pacific                  |    49
 Hispanic                       |   237
 Black                          |   296
 White                          |  1158
(5 rows)
```

Total: 1745

Percentage of officers of each identity group with allegations against them (compared to the total) is written in bold type.

Native American: 0.002865329513 * 100 = **0.2865%** (0.33%)

Asian/Pacific: 0.02808022923 * 100 = **2.808%** (11.59%)

Hispanic: 0.1358166189 * 100 = **13.5817%** (n/a)

Black: 0.1696275072 * 100 = **16.9628%** (8.94%)

White: 0.6636103152 * 100 = **66.3610%** (75.49%)

b) The percentage of that identity group in the general population is written in parentheses in gray next to the part a percentages

2) Lincoln Park: Avg Income $71,551.00
   a) Table of number of allegations per identity group in this neighborhood

```
SELECT race, COUNT(*) FROM data_oa_with_area WHERE name='1813' OR
name='1812' OR name='1811' OR name='1932' OR name='1935' OR
name='1814' GROUP BY race ORDER BY COUNT(*) ASC;
```

```
     race      | count
---------------+-------
 Asian/Pacific |     1
 Hispanic      |    31
 Black         |    66
 White         |   200
(4 rows)
```

Total: 298

Percentage of officers of each identity group with allegations against them (compared to the total) is written in bold type.

Asian/Pacific: 0.003355704698 * 100 = **0.33557%** (6.18%)

Hispanic: 0.1040268456 * 100 = **10.4027%** (n/a)

Black: 0.2214765101 * 100 = **22.1477%** (4.69%)

White: 0.6711409396 * 100 = **67.1141%** (85.23%)

b) The percentage of that identity group in the general population is written in parentheses in gray next to the part a percentages

3) Lake View: Avg Income $60,058.00
   a) Table of number of allegations per identity group in this neighborhood

```
SELECT race, COUNT(*) FROM data_oa_with_area WHERE name='1935' OR
name='1934' OR name='1933' OR name='1932' OR name='1924' OR
name='1923' OR name='1922' GROUP BY race ORDER BY COUNT(*) ASC;
     race      | count
---------------+-------
 Asian/Pacific |     4
 Black         |    14
 Hispanic      |    27
 White         |   114
(4 rows)
```

Total: 159

Percentage of officers of each identity group with allegations against them (compared to the total) is written in bold type.

Asian/Pacific: 0.0251572327 * 100 = **2.5157%** (6.77%)

Hispanic: 0.1698113208 * 100 = **16.9811%** (n/a)

Black: 0.08805031447 * 100 = **8.805%** (3.40%)

White: 0.7169811321 * 100 = **71.6981%** (85.50%)

b) The percentage of that identity group in the general population is written in parentheses in gray next to the part a percentages

Lower Income neighborhoods (descending order):

1) North Lawndale: Avg. Income $12,034.00
   a) Table of number of allegations per identity group in this neighborhood

```
SELECT race, COUNT(*) FROM data_oa_with_area WHERE name='1133' OR
name='1023' OR name='1022' OR name='1021' OR name='1014' OR
name='1013' OR name='1012' OR name='1011' GROUP BY race ORDER BY
COUNT(*) ASC;
              race              | count
--------------------------------+-------
 Native American/Alaskan Native |     8
 Asian/Pacific                  |    58
```

```
 Black                            |   291
 Hispanic                         |   993
 White                            |  1199
(5 rows)
```

Total: 2,549

Percentage of officers of each identity group with allegations against them (compared to the total) is written in bold type.

Native American: 0.003138485681 * 100 = **0.3138%** (0.25%)

Asian/Pacific: 0.02275402118 * 100 = **2.2754%** (0.25%)

Hispanic: 0.3895645351 * 100 = **38.9564%** (n/a)

Black: 0.1141624166 * 100 = **11.4162%** (88.47%)

White: 0.4703805414 * 100 = **47.0381%** (7.34%)

c) The percentage of that identity group in the general population is written in parentheses in gray next to the part a percentages

2) West Garfield Park: Avg. Income $10,934.00

a) Table of number of allegations per identity group in this neighborhood

```
SELECT race, COUNT(*) FROM data_oa_with_area WHERE name='1123' OR
name='1124' OR name='1134' OR name='1222' OR name='1124' OR
name='1125' OR name='1133' OR name='1122' OR name='1132' OR
name='114' OR name='115' GROUP BY race ORDER BY COUNT(*) ASC;
              race              | count
--------------------------------+-------
 Native American/Alaskan Native |    19
 Asian/Pacific                  |    66
 Black                          |   615
 Hispanic                       |   835
 White                          |  1993
(5 rows)
```

Total: 3528

Percentage of officers of each identity group with allegations against them (compared to the total) is written in bold type.

Native American: 0.005385487528 * 100 = **0.5385%** (0.13%)

Asian/Pacific: 0.01870748299 * 100 = **1.8707%** (0.19%)

Hispanic: 0.2366780045 * 100 = **23.6678%** (n/a)

Black: 0.1743197279 * 100 = **17.43197%** (95.58%)

White: 0.5649092971 * 100 = **56.4909%** (2.16%)

d) The percentage of that identity group in the general population is written in parentheses in gray next to the part a percentages

3) South Lawndale: Avg. Income $10,402.00
    a) Table of number of allegations per identity group in this neighborhood

```
SELECT race, COUNT(*) FROM data_oa_with_area WHERE name='1013' OR
name='1024' OR name='1031' OR name='1032' OR name='1033' GROUP BY
race ORDER BY COUNT(*) ASC;
```

race	count
Native American/Alaskan Native	3
Asian/Pacific	10
Black	88
White	225
Hispanic	288

Total: 614

Percentage of officers of each identity group with allegations against them (compared to the total) is written in bold type.

Native American: 0.004885993485 * 100 = **0.4886%** (0.54%)

Asian/Pacific: 0.01628664495 * 100 = **1.6287%** (6.09%)

Hispanic: 0.4690553746 * 100 = **46.9055% (n/a)**

Black: 0.1433224756 * 100 = **14.3322%** (30.95%)

White: 0.3664495114 * 100 = **36.64495%** (48.69%)

e) The percentage of that identity group in the general population is written in parentheses in gray next to the part a percentages

WHAT WE LEARNED FROM QUESTION 3:

In higher income neighborhoods, Black officers were *more* likely to have allegations against them compared to the overall population of black people while white officers were *less* likely to have allegations against them compared to the general population in that neighborhood. Interestingly, Asian officers were *much* less likely to have allegations against them compared to the overall population, but this could be due simply to the small sample size. In lower income neighborhoods, white officers were much more likely to have allegations against them compared to the general population while black officers were much less likely to have allegations against them compared to the general population. From this, we learned that income area matters a lot in correlating whether an officer of a certain race is more or less likely to have an allegation against them compared to the general population.

4. Are certain identity groups overrepresented as complainants in low income areas verses higher income areas compared to the general population?

We used Trifacta Wrangler to join the data_complainants table with information from data_allegation and data_area to combine demographic information about the people who filed complaints with the areas (beats) that the allegations took place.

Queries to import CSV to Postgres:

```
CREATE TABLE data_complainant_with_area (beat_id varchar, id varchar,
allegation_id varchar, id1 varchar, id2 varchar, gender varchar, race
varchar, age varchar, birth_year varchar, name varchar);

\copy data_complainant_with_area FROM
'/Users/caseygrage/Downloads/data_complainant_with_area.csv' WITH
(FORMAT csv);
```

Used https://www.chicagocomputerclasses.com/average-city-chicago-income/ to find the neighborhoods with the highest and lowest average incomes.

Used https://data.cpdp.co/data/bgKwNB/ to find the beats within neighborhoods of interest.

Used www.areavibes.com to find racial demographics of neighborhoods of interest.

Richest neighborhoods (descending order):

4) Near North Side: Avg Income $88,669.00
    a) Table of number of allegations per identity group in this neighborhood

```
SELECT race, COUNT(*) FROM data_complainant_with_area WHERE
name='1833' OR name='1834' OR name='1832' OR name='1831' OR
name='1824' OR name='1823' OR name='1822' OR name='1821' GROUP BY
race ORDER BY COUNT(*) ASC;
```

race	count
Native American/Alaskan Native	8
Asian/Pacific Islander	37
Hispanic	150
White	841
Black	1281

Total: 2317

Percentage of complainants of each identity group (compared to the total) is written in bold type.

Native American/Alaskan Native: .003452741 * 100 = **0.3453%** (0.33%)

Asian/Pacific Islander: .015968925 * 100 = **1.5970%** (11.59%)

Hispanic: .064738886 * 100 = **6.4739%** (n/a)

White: .362969357 * 100 = **36.2969%** (75.49%)

Black: .552870091 * 100 = **55.287%** (8.94%)

b) The percentage of that identity group in the general population is written in parentheses in gray next to the part a percentages

5) Lincoln Park: Avg Income $71,551.00

a) Table of number of allegations per identity group in this neighborhood

```
SELECT race, COUNT(*) FROM data_complainant_with_area WHERE
name='1813' OR name='1812' OR name='1811' OR name='1932' OR
name='1935' OR name='1814' GROUP BY race ORDER BY COUNT(*) ASC;
```

```
          race          | count
------------------------+-------
 Asian/Pacific Islander |    10
 Hispanic               |    69
 Black                  |   168
 White                  |   216
(4 rows)
```

Total: 463

Percentage of complainants of each identity group (compared to the total) is written in bold type.

Asian/Pacific Islander: .021598272 * 100 = **2.1598%** (6.18%)

Hispanic: .149028078 * 100 = **14.9028%** (n/a)

Black = .362850972 * 100 = **36.2851%** (4.69%)

White = .466522678 * 100 = **46.6523%** (85.23%)

From https://www.areavibes.com/chicago-il/lincoln+park/demographics/

b) The percentage of that identity group in the general population is written in parentheses in gray next to the part a percentages

6) Loop: Avg Income $65,526.00

a) Table of number of allegations per identity group in this neighborhood

```
SELECT gender, race, COUNT(*) FROM data_complainant_with_area WHERE
name='112' OR name='114' OR name='111' OR name='122' OR name='123'
GROUP BY gender, race ORDER BY COUNT(*) ASC;
 gender | race | count
--------+------+-------
(0 rows)
```

NO DATA

7) Lake View: Avg Income $60,058.00
    a) Table of number of allegations per identity group in this neighborhood

```
SELECT race, COUNT(*) FROM data_complainant_with_area WHERE
name='1935' OR name='1934' OR name='1933' OR name='1932' OR
name='1924' OR name='1923' OR name='1922' GROUP BY race ORDER BY
COUNT(*) ASC;
              race              | count
--------------------------------+-------
 Native American/Alaskan Native |     4
 Asian/Pacific Islander         |    20
 Hispanic                       |   116
 Black                          |   320
 White                          |   468
```

Total: 928

Percentage of complainants of each identity group (compared to the total) is written in bold type.

Native American/Alaskan Native: .004310345 * 100 = **0.431%** (0.27%)

Asian/Pacific Islander: .021551724 * 100 = **2.1552%** (6.77%)

Hispanic: .125 * 100 = **12.5%** (n/a)

Black: .344827586 * 100 = **34.4828%** (3.40%)

White: .504310345 * 100 = **50.431%** (85.50%)

    b) The percentage of that identity group in the general population is written in parentheses in gray next to the part a percentages

3 Poorest Neighborhoods (descending order)

4) North Lawndale: Avg. Income $12,034.00
    a) Table of number of allegations per identity group in this neighborhood

```
SELECT race, COUNT(*) FROM data_complainant_with_area WHERE
name='1133' OR name='1023' OR name='1022' OR name='1021' OR
```

```
name='1014' OR name='1013' OR name='1012' OR name='1011' GROUP BY
race ORDER BY COUNT(*) ASC;
          race          | count
------------------------+-------
 Asian/Pacific Islander |     9
 White                  |   158
 Hispanic               |   199
 Black                  |  1969
(4 rows)
```

Total: 2335

Percentage of complainants of each identity group (compared to the total) is written in bold type.

Asian/Pacific Islander: .00385439 * 100 = **0.3854%** (0.25%)

White: .067665953 * 100 = **6.7666%** (7.34%)

Hispanic: .085224839 * 100 = **8.5225%** (n/a)

Black: .843254818 * 100 = **84.3255%** (88.47%)

b) The percentage of that identity group in the general population is written in parentheses in gray next to the part a percentages

5) Englewood: Avg. Income $11,888.00
   a) Table of number of allegations per identity group in this neighborhood

```
SELECT race, COUNT(*) FROM data_complainant_with_area WHERE
name='734' OR name='724' OR name='713' OR name='725' OR name='715' OR
name='714' OR name='726' OR name='735' GROUP BY race ORDER BY
COUNT(*) ASC;
 race | count
------+-------
(0 rows)
```

NO DATA

6) West Englewood: Avg. Income $11,317.00
   a) Table of number of allegations per identity group in this neighborhood

```
SELECT race, COUNT(*) FROM data_complainant_with_area WHERE
name='734' OR name='724' OR name='713' OR name='725' OR name='715' OR
name='714' OR name='726' OR name='735' OR name='712' OR name='711' OR
name='722' OR name='731' OR name='732' GROUP BY race ORDER BY
COUNT(*) ASC;
 race | count
```

```
------+-------
(0 rows)
```

NO DATA

7) West Garfield Park: Avg. Income $10,934.00
    a) Table of number of allegations per identity group in this neighborhood

```
SELECT race, COUNT(*) FROM data_complainant_with_area WHERE
name='1123' OR name='1124' OR name='1134' OR name='1222' OR
name='1124' OR name='1125' OR name='1133' OR name='1122' OR
name='1132' OR name='114' OR name='115' GROUP BY race ORDER BY
COUNT(*) ASC;
              race              | count
--------------------------------+-------
 Native American/Alaskan Native |     7
 Asian/Pacific Islander         |    16
 Hispanic                       |   153
 White                          |   559
 Black                          |  2496
```

Total: 3231

Percentage of complainants of each identity group (compared to the total) is written in bold type.

Native American/Alaskan Native: .002166512 * 100 = **0.2167%** (0.13%)

Asian/Pacific Islander: .004952027 * 100 = **0.4952%** (0.19%)

Hispanic: .04735376 * 100 = **4.7354%** (n/a)

White: .173011452 * 100 = **17.3011%** (2.16%)

Black: .772516249 * 100 = **77.2516%** (95.58%)

    b) The percentage of that identity group in the general population is written in parentheses in gray next to the part a percentages

8) Fuller Park: Avg Income $10,432.00
    a) Table of number of allegations per identity group in this neighborhood

```
SELECT gender, race, COUNT(*) FROM data_complainant_with_area WHERE
name='225' OR name='224' OR name='213' OR name='215' OR name='935' OR
name='925' GROUP BY gender, race ORDER BY COUNT(*) ASC;
 gender | race | count
--------+------+-------
(0 rows)
```

NO DATA

9) South Lawndale: Avg. Income $10,402.00
    a) Table of number of allegations per identity group in this neighborhood

QUERY:
```
SELECT gender, race, COUNT(*) FROM data_complainant_with_area
WHERE name='1013' OR name='1024' OR name='1031' OR name='1032' OR
name='1033' GROUP BY gender, race ORDER BY COUNT(*) ASC;
```

```
 gender |          race          | count
--------+------------------------+-------
 M      | Asian/Pacific Islander |     2
 F      | Asian/Pacific Islander |     2
 F      | White                  |    75
 M      | White                  |   109
 F      | Black                  |   173
 M      | Black                  |   198
 F      | Hispanic               |   261
 M      | Hispanic               |   415
(8 rows)
```

Total: 1235

Percentage of complainants of each identity group (compared to the total) is written in bold type.

Asian/Pacific Islander: .003238866 * 100 = **0.3239%** (6.09%)

White: .148987854 * 100 = **14.8988%** (48.69%)

Black: .300404858 * 100 = **30.0405%** (30.95%)

Hispanic: .547368421 * 100 = **54.7368%** (n/a)

    b) The percentage of that identity group in the general population is written in parentheses in gray next to the part a percentages

10) Riverdale: Avg Income $8,201.00
    a) Table of number of allegations per identity group in this neighborhood

```
SELECT gender, race, COUNT(*) FROM data_complainant_with_area WHERE
name='533' OR name='532' GROUP BY gender, race ORDER BY COUNT(*) ASC;
```

```
 gender | race | count
--------+------+-------
(0 rows)
```

NO DATA

WHAT WE LEARNED FROM QUESTION 4:

In high income neighborhoods, Black civilians were disproportionately *more likely* to file complaints against officers than their reflected populations in the area, in many cases on a scale of 5x to 10x higher than their percentage of the population. On the other hand, White civilians were disproportionately *less likely* to file complaints in proportion to their overall percentage of the population. This may suggest underlying differences in treatment by police officers towards civilians of different races. In lower income areas, this disparity is not as apparent. It is interesting to note that the demographic makeups of these areas are not as "white-dominant," so it may be difficult to draw wide conclusions among lower-income areas. Despite this, in general, Black civilians remain the most likely group to file complaints against officers, suggesting a common pattern.